AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions and listings of claims in the application:

LISTING OF CLAIMS:

1. (currently amended) A method for organizing a plurality of data files using meta-data wherein at least one meta-data element is associated with each data file, the method comprising:

extracting~~, for at least some selected data files,~~ the at least one meta-data element associated with ~~a respective selected~~ each data file;

organizing the extracted meta-data elements into a single ordered set wherein the set is ordered consecutively based on values ~~for one or more~~ of the extracted meta-data elements;

calculating pair-wise differences between values ~~of each~~ of the extracted meta-data elements for all possible pairs of data files;

inputting at least one value of a clustering sensitivity parameter, said clustering sensitivity parameter defining granularity of the clustering, and multiplying each pair-wise difference by the clustering sensitivity parameter to obtain a plurality of similarity values for determining clustering based on the similarity values and at a granularity defined by the clustering sensitivity parameter; and

dividing the selected data files into groups based on the similarity values.

2. canceled.

3. (previously presented): The method of claim 1, wherein determining the similarity value of a one pair comprises evaluating the similarity value as:

$$S_K(i,j) = \exp\left(-\frac{|t_i - t_j|}{K}\right),$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value;

t_i is the value of a selected extracted meta-data element associated with the i^{th} data file; and

t_j is the value of a selected extracted meta-data element associated with the j^{th} data file.

4. (previously presented): The method of claim 1, wherein determining the similarity value of a one pair comprises determining the similarity value of the one pair as:

$$S_K(i,j) = \exp\left(\frac{1}{K}\left(-\frac{< v_i, v_j >}{|v_i||v_j|} - 1\right)\right).$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value; and

v_i and v_j are actual vector values determined from the i^{th} and the j^{th} data files.

5. (previously presented): The method of claim 1, further comprising determining, for each of the selected data files, at least one novelty value based on at least one similarity value of the selected data file and a nearby data file.

6. (currently amended): The method of claim 5, wherein determining at least one novelty value comprises determining at least one novelty value as:

$$[[v_K(s) = \sum_{l,n} S_K(s+1, s+n) g(1,n)]]$$

$$v_K(s) = \sum_{l,n} S_K(s+l, s+n) g(l,n)$$

where:

$v_K(s)$ is the novelty value; and

g is a Gaussian tapered *(2l+1) x (2n+1)* checkerboard kernel.

7. (previously presented): The method of claim 5, further comprising determining at least one boundary location between the selected data files based on the at least one novelty values determined for each of the selected data files.

8. (previously presented): The method of claim 7, further comprising determining, for at least the one determined boundary locations, a confidence value for that boundary location.

9. (previously presented): The method of claim 8, wherein determining a confidence value for a boundary location comprises determining the confidence value as:

$$C(B_K) = \sum_{l=1}^{|B_K|-1} \frac{1}{(b_{l+1} - b_1)^2} \sum_{i,j=b1}^{b_{l+1}} S_K(i,j) - \sum_{l=1}^{|B_K|-2} \frac{1}{(b_{l+1} - b_1)(b_{l+2} - b_{l+1})} \sum_{i=b_1}^{b_{l+1}} \sum_{j=b_{l+1}}^{b_{l+2}} S_K(i,j) \ .$$

where:

$C(B_K)$ is the confidence value for the B_K^{th} boundary;

$S_K(i, j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

b is the index value of detected boundary at a particular value for the input clustering sensitivity parameter K level.

10. (previously presented): The method of claim 8, further comprising determining, for at least the one determined boundary location~~s~~, at least one clustering sensitivity parameter value that maximizes the confidence value.

11. (currently amended): A method for organizing a plurality of data files stored in a digital memory using meta-data, wherein at least one meta-data element is ~~at least~~ associated with a corresponding one of the plurality of data files, the method comprising:

extracting from the memory meta-data elements of the plurality of data files;

organizing the extracted meta-data elements into a single ordered set wherein the set is ordered consecutively based on values for ~~one or more of~~ the extracted meta-data elements;

calculating pair-wise difference between values ~~of each~~ of the extracted meta-data elements for all possible pairs of data files;

obtaining a desired value of a clustering sensitivity parameter for analyzing the meta-data;

multiplying each pair-wise difference by the clustering sensitivity parameter to obtain a plurality of similarity values having granularity defined by the clustering sensitivity parameter;

determining a structure within the meta-data elements by comparing, for at least a subset of the plurality of data files, the similarity values; and [[,]]

storing the structure of the data files in a memory.

12. (previously presented): The method of claim 11, further comprising clustering the data files into groups using the determined structure of the meta-data elements.

13. (original): The method of claim 12, further comprising determining boundaries from the determined clusters of data files, wherein the boundaries are located between the determined clusters of data files.

14. (previously presented): The method of claim 13, further comprising:

determining a similarity value by comparing at least some of the meta-data elements associated with data files in a first selected cluster of data files to at least some other meta-data elements associated with data files of in the first selected cluster of data files; and

determining a dissimilarity value by comparing at least some of the meta-data elements associated with data files in a second selected cluster of data files to at least some of the meta-data elements associated with data files in a third selected cluster of data files, wherein the second selected cluster of data files and the third selected cluster of data files are different clusters of data files.

15. (original): The method of claim 14, further comprising:

determining a value corresponding to a desired grouping of the clusters of data files based on the differences of the similarity values and the dissimilarity values.

16. (currently amended): A storage medium storing a set of program instructions executable on a data processing device and usable to organize a plurality of data files by using meta-data wherein at least one meta-data element is associated with each data file, the program comprising:

instructions for extracting~~, for at least some of the data files,~~ the at least one meta-data element associated with each of the ~~some~~ data files;

instructions for organizing the extracted meta-data elements into a single ordered set wherein the set is ordered consecutively based on at least one of: chronological, alphabetical, numerical, or geographical ordering;

instructions for calculating pair-wise difference between values of ~~each of~~ the extracted meta-data elements for all possible pairs of data files;

instructions for inputting at least one clustering sensitivity parameter value, said clustering sensitivity parameter defining granularity of the clustering;

instructions for multiplying each pair-wise difference by the clustering sensitivity parameter to obtain a plurality of similarity values having granularity defined by the clustering sensitivity parameter; and

instructions for dividing the data files into groups based on the similarity values of the extracted meta-data elements.

17. canceled.

18. (previously presented): The storage medium of claim 16, further comprising instructions for determining, for each of the at least some data files, at least one novelty value for that data file based on at least one similarity value of that data file and a nearby data file.

19. (previously presented): The storage medium of claim 16, wherein instructions for determining the similarity value of the one pair comprises instructions for evaluating the similarity value as:

$$S_K(i,j) = \exp\left(-\frac{|t_i - t_j|}{K}\right),$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value;

t_i is the value of a selected extracted meta-data element associated with the i^{th} data file; and

t_j is the value of a selected extracted meta-data element associated with the j^{th} data file.

20. (previously presented): The storage medium of claim 16, wherein instructions for determining the similarity value of the one pair comprises determining the similarity value of the one pair as:

$$S_K(i,j) = \exp\left(\frac{1}{K}\left(-\frac{< v_i - v_j >}{|v_i||v_j|} - 1\right)\right).$$

where:

the one pair consists of the i^{th} data file and j^{th} data file;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

K is the clustering sensitivity parameter value; and

v_i and v_j are actual vector values determined from the i^{th} and the j^{th} data files.

21. (previously presented): The storage medium of claim 18, further comprising instructions for determining at least one boundary location between the some data files based on the at least one novelty values determined for the each of the at least some data files.

22. (currently amended): The storage medium of claim 18, wherein instructions for determining at least one novelty value comprises instructions for determining the at least one novelty value as:

$$[[v_K(s) = \sum_{l,n} S_K(s+1, s+n)g(1,n)]]$$

$$v_K(s) = \sum_{l,n} S_K(s+l, s+n)g(l,n)$$

where:

$v_K(s)$ is the novelty value; and

g is the Gaussian tapered $(2l+1) \times (2n+1)$ checkerboard kernel.

23. (previously presented): The storage medium of claim 21, further comprising instructions for determining, for at least the one determined boundary locations, a confidence value for that boundary location.

24. (original): The storage medium of claim 23, wherein instructions for determining at least one confidence value comprises instructions for determining each of such confidence value as:

$$C(B_K) = \sum_{l=1}^{|B_K|-1} \frac{1}{(b_{l+1} - b_1)^2} \sum_{i,j=b1}^{b_{l+1}} S_K(i,j) - \sum_{l=1}^{|B_K|-2} \frac{1}{(b_{l+1} - b_1)(b_{l+2} - b_{l+1})} \sum_{i=b_1}^{b_{l+1}} \sum_{j=b_{l+1}}^{b_{l+2}} S_K(i,j) \ .$$

where:

$C(B_K)$ is the confidence value for the B_K^{th} boundary;

$S_K(i,j)$ is the similarity value for the i^{th} data file and the j^{th} data file;

b is the detected boundary at a level.

25. (previously presented): The storage medium of claim 23, further comprising instructions for determining, for at least the one determined boundary location, at least one clustering sensitivity parameter value that maximizes the confidence value.

26. (currently amended): A computerized system usable to organize a plurality of data files using meta data having at least one meta data element that is ~~at least~~ associated with a corresponding one of the data files, comprising:

a memory for storing the data files;

an input link for receiving the data files;

a meta-data extracting circuit [[, routine, or application]] that extracts, for ~~at least some of~~ the data files, the at least one meta-data element associated with that data file;

a meta-data organizing circuit [[, routine or application]] that organizes the extracted meta-data elements in a desired order based on values for the extracted meta-data elements;

a similarity value determining circuit [[, routine or application]] that:

selects at least one clustering sensitivity parameter value, said clustering sensitivity parameter defining granularity of clustering [[;]] , and

calculates a similarity value for ~~at least two of~~ all possible pairs of data files in the plurality of data files by performing a pair-wise comparison of the extracted meta-data

elements, corresponding to all possible pairs of data files, and multiplying the pair-wise comparison by the clustering sensitivity parameter value;

a novelty value determining circuit [[, routine or application]] that determines at least one novelty value for that data file based on the at least one similarity value for that data file and for a number of nearby data files;

a data dividing determining circuit [[, routine or application]] that divides at least some of the data files into groups based on the extracted meta-data elements and the input clustering sensitivity parameter value by determining at least one boundary location between ones of the plurality of data files based on the at least one novelty value determined for at least some of the data files; and

a confidence value determining circuit [[, routine or application]] that determines, for at least some of the determined boundary locations, a confidence value for that boundary location, wherein the data dividing circuit [[, routine, or application]] further determines, for at least some of the determined boundary locations, the at least one clustering sensitivity parameter value that maximizes the confidence value; and [[,]]

an output link for outputting the meta-data elements in a desired grouping order,

wherein one or more processors control operation of the circuits.

27. (previously presented): The method of claim 1, wherein determining the similarity value of the one pair comprises a term depending on a scalar magnitude of a difference

between t_i and t_j relative to K, where K is the clustering sensitivity parameter value, i references a first data file of the one pair, j references a second file of the one pair, t_i is the value of a selected extracted meta-data element associated with the first data file, and t_j is the value of a selected extracted meta-data element associated with the second data file.

28. (previously presented): The method of claim 27 wherein the depending comprises an exponentially decreasing function of the scalar magnitude of the difference between t_i and t_j relative to K.

29. (previously presented): The method of claim 1, wherein determining the similarity value of the at least one pair of the selected data files comprises a term depending on an inner product of v_i and v_j relative to K, where K is the clustering sensitivity parameter value, v_i is an actual vector value determined from the i^{th} data file, and v_j is an actual vector value determined from the j^{th} data file.